



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Understanding Equity, Diversity and Inclusivity Challenges Within the Research Software Community

Citation for published version:

Chue Hong, NP, Cohen, J & Jay, C 2021, Understanding Equity, Diversity and Inclusivity Challenges Within the Research Software Community. in *Computational Science – ICCS 2021*. vol. VI, Lecture Notes in Computer Science, vol. 12747, Springer, pp. 390–403, International Conference on Computational Science 2021, Krakow, Poland, 16/06/21. https://doi.org/10.1007/978-3-030-77980-1_30

Digital Object Identifier (DOI):

https://doi.org/10.1007/978-3-030-77980-1_30

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Computational Science – ICCS 2021

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Understanding Equity, Diversity and Inclusion Challenges Within the Research Software Community

Neil P. Chue Hong¹[0000–0002–8876–7606], Jeremy Cohen²[0000–0003–4312–2537],
and Caroline Jay³[0000–0002–6080–1382]

¹ Software Sustainability Institute & EPCC, University of Edinburgh, Edinburgh, U.K.

N.ChueHong@epcc.ed.ac.uk

² Department of Computing, Imperial College London, London, U.K.

jeremy.cohen@imperial.ac.uk

³ Department of Computer Science, University of Manchester, Manchester, U.K.

Caroline.Jay@manchester.ac.uk

Abstract. Research software – specialist software used to support or undertake research – is of huge importance to researchers. It contributes to significant advances in the wider world and requires collaboration between people with diverse skills and backgrounds. Analysis of recent survey data provides evidence for a lack of diversity in the Research Software Engineer community. We identify interventions which could address challenges in the wider research software community and highlight areas where the community is becoming more diverse. There are also lessons that are applicable, more generally, to the field of software development around recruitment from other disciplines and the importance of welcoming communities.

Keywords: research software, software engineering, research software engineering, diversity, EDI.

1 Introduction

Developing specialist research software to support computational science is an especially challenging process. Unlike more traditional software engineering tasks, researchers or, increasingly, RSEs (Research Software Engineers) who write this software need an understanding of the underlying scientific challenge being addressed.

The methods used in research favour a hypothesis-driven approach. This creates a different working environment from the wider software engineering industry, where software is built to meet a client’s specification. As noted by Hettrick et al.,

“To be effective, software development in research should be approached, not as a one-off transaction, but as a partnership between researcher and software expert.” [3]

The “partnership between researcher and software expert” mentioned here highlights a need for varied skills and good communication. It is clear that much computational research requires diversity of skills and experience, working in partnership, that spans software engineering and science. Researchers are more likely to be working with others who have different technical expertise, use different technical terminology, and may be communicating in a tertiary language.

Diversity in project teams, and workplaces in general, is important and it is widely accepted that there are many benefits from ensuring diversity within teams and communities. For example, there is evidence to show that diversity in terms of knowledge or skills can be beneficial [19] and that gender diversity within groups can result in higher quality scientific outputs [4]. However, there are cases where diversity can raise challenges, for example with reduced feelings of well-being amongst members in highly-diverse teams [31]. Diversity in the context of teams in industry and research is an area that has been the subject of extensive research and there are some contradictory results from the many studies undertaken. Stahl et al. [26] look at a number of different studies of cultural diversity concluding that there are both benefits and drawbacks. Whether benefits can be realised while minimising more challenging aspects will depend on effective process management. In the context of software engineering, Capretz & Ahmed [5] looked at how different personality traits relate to suitability for different roles in software projects. Ultimately they conclude that diversity in terms of personalities and skills is important in helping with problem solving tasks involved in building and maintaining software.

The different types of diversity highlighted so far, including diversity of skills and knowledge, culture, personality and gender are, of course, just some of the many different aspects that lead to diversity amongst a group of individuals. Others, among a huge range, include ethnicity, disability and age.

Research Software Engineers are, at present, much more likely to come from a research background than software engineering professionals working in other fields. Nonetheless, we know from surveys within the RSE community that a similar diversity crisis to that identified in other fields exists. In this paper, we examine current problems with diversity in research software engineering and consider potential causes. Our aim is to present a better understanding of existing diversity within the RSE community in order to develop insights and recommendations to address current issues. We achieve this through an empirical analysis of existing open data collected and made available through large-scale international surveys of research software engineers undertaken by the RSE community. We also review related work from allied areas which suggests ways of addressing the challenges identified.

As pointed out by Mathieson in [20], based on 2018 ONS data, just one in eight of almost 340,000 software development professionals in the UK are women. This example demonstrates that there is a long way to go to address the lack of diversity in the wider software engineering domain, as well as the more focused domain of research software engineers. Professional software engineers can gain their skills through a variety of different routes and they may come from a wide

range of different disciplines. For example, they may have developed their skills through a degree programme or vocational training in computing or a related area. Alternatively, they may come from a completely different disciplinary background and have re-trained through one of a large number of “coding schools” that offer training, often via intensive courses, in software development skills. The wide array of routes into the field creates an expectation of a diverse field which makes it all the more surprising that this is not the case.

We highlight three core contributions this paper provides to understanding the importance of diversity among RSEs working in computational and data science:

- Providing an evidence-based analysis that demonstrates the current problems with diversity in the research software engineering community.
- Demonstrating that there is already extensive “domain mobility” for RSEs and that a lack of such mobility is therefore not likely to be a cause for a lack of diversity within the field.
- Offering four general recommendations that we believe can form a basis to support addressing the lack of diversity in research software engineering.

In Section 2 we examine the International RSE Survey results from 2018 and undertake further analysis on this data. Section 3 highlights the three areas that we see as both helping to explain and provide the basis for addressing the lack of diversity amongst RSEs, while discussion and conclusions are provided in Section 4.

2 Surveying RSE Diversity

To understand the diversity challenges facing those embarking on careers as Research Software Engineers, we require a better understanding of the landscape as a whole. The Software Sustainability Institute coordinates an ongoing series of surveys of the RSE community. In the most recent survey [24], from 2018, participants were asked a number of socio-demographic questions.

Percentage of UK who are:	RSEs[24]	Academics[15]	Software Developers[25]	All UK workers[25]
Gender (female)	14	46	14	48
Ethnicity (BAME/Mixed)	5	15	21	12
Report disability	6	4	10	13

Table 1. Comparison of 2018 demographics for Research Software Engineers, Academics, Software Developers and general working population in the United Kingdom.

Our reanalysis primarily focuses on UK data, as this is where the authors are located, but the numbers from other countries in the survey are broadly similar. Table 1 compares RSEs with UK Higher Education Statistics Authority (HESA) data [15], and a study by the British Computer Society (BCS) of software development professionals in the UK [25] based on 2018 Office for National

Statistics data. Limitations in the RSE data currently available to us mean that we focus on aspects of gender, ethnicity and disability, as opposed to other types of diversity.

This comparison indicates a gender diversity gap, which we might expect given the percentage of women working as software developers is also 14%. But this could be better: software is a fundamental part of all research, and 46% of academic staff are female. The ethnicity data indicates a greater problem. Of the respondents who declared their ethnicity, there were only 6 non-white and 5 mixed race RSEs – 5%. These figures are significantly lower than the 28% BAME students studying computer science [30] and 10% of BAME physicists [17]. Although we should not aggregate non-white ethnic categories together, as they will experience different challenges and biases, it is striking that RSEs do not fit the general profile of those working in the IT industry, where there is greater ethnic diversity.

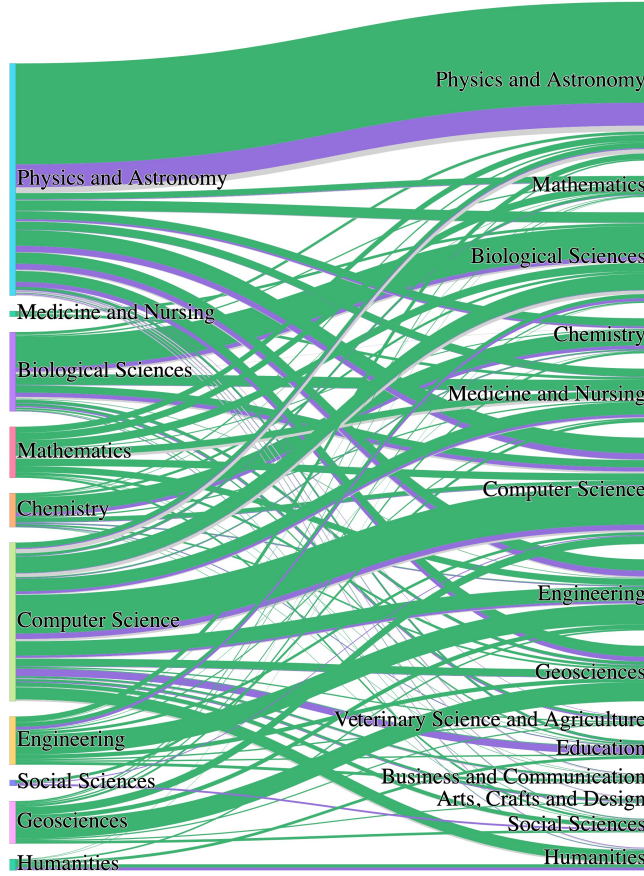


Fig. 1. Flow of UK RSEs from academic discipline studied to domains they now work in, categorised by gender: 79.8% male (green) / 14.3% female (purple) / 5.9% other responses (grey). N=203.

An important area of research is understanding if diversity can be improved if RSEs are drawn from a wider set of degree backgrounds. Figure 1 shows where RSEs are working based on their academic degree. Only gender diversity was considered, given the low ethnic diversity and number of respondents reporting a disability. Gender balance is relatively uniform across fields rather than mirroring the gender balance in that field. Over half of RSEs have a first degree in Physics and Astronomy or Computer Science. In the UK, 17% of CS undergraduates were female compared with 41% of physical sciences undergraduates [16], suggesting that RSEs come from the “computational” subset of a subject. However, within computer science research within the UK just under 23% of academics and researchers are female [14]. This perhaps suggests that a larger percentage of female CS undergraduates move on to a research or faculty position in CS than male undergraduates, or that female CS researchers are moving into the field after undergraduate studies in other areas. However, it could also be attributed to other factors such as industry hiring trends. Of interest here is that, while over half of RSEs have a Physics and Astronomy or Computer Science-focused first degree, the percentage of female RSEs is still somewhat below the percentage of female CS research/faculty staff.

The areas that RSEs support are more widespread than the fields they are currently recruited from. There are teams offering general RSE support at a number of institutions and many RSEs working in the biosciences, geosciences, and medicine. This suggests that a challenge research software engineering has is encouraging more people with first degrees in these areas (which have more balanced gender representation) to become RSEs. This is important because domain knowledge can be especially valuable when undertaking RSE projects. Things to consider include the wording of job adverts, which can affect who applies for them [1] — can we encourage more candidates from other disciplines? There may be lessons for the software development field more generally — there are a significant number of RSEs whose background is in biological sciences or geosciences, as well as some from social sciences and the humanities, showing people with “non-traditional” degree backgrounds seek careers as software practitioners.

One encouraging aspect of Figure 1 is there appear to be many lines going from one discipline on the left hand side to a different discipline on the right hand side, representing an RSE working in an area different from the one where they trained. This suggests that there is scope to improve diversity within the RSE community through attracting individuals from more diverse domains into the RSE space. This provides an opportunity for more immediate improvements in the diversity of the RSE community than the multi-year timeline that we might expect for efforts to improve diversity amongst undergraduate cohorts. To investigate this movement between domains, which we call “*domain mobility*”, we undertook further analysis of the most recent 2018 RSE survey data.

Table 2 shows an example of “domain mobility” for a series of domains where the number of survey respondents was > 20 . We define *partial mobility* as RSEs working in one or more domains outside the domain of their highest degree, while

Discipline	N	Partially mobile (%)	Fully mobile (%)
Biological Sciences	95	44.21	8.42
Chemistry	22	40.91	22.73
Computer Science	241	51.45	22.41
Electrical & Electronic Engineering	36	41.67	38.89
Geography & Environmental Sciences	46	52.17	6.52
Mathematics	69	47.83	47.83
Physics and Astronomy	261	31.80	19.92

Table 2. Example of domain mobility across a selection of disciplines from the 2018 RSE survey data where $N > 20$.

also still undertaking work in the domain of their degree. *Full mobility* applies a further filter to partial mobility by excluding RSEs who still work in the domain of their highest degree. So, in the case of Mathematics, where the figures for partial and full mobility are the same, none of the survey respondents who said their highest degree was in Mathematics undertake RSE work in the domain of Mathematics. It should be noted that while domain mobility is a positive concept, and the ability to work across a wide variety of different domains is a great benefit, we do not have any data on the reasons for these movements between domains. The movements may be individuals wanting to undertake RSE work in, and learn about, a different domain to their original area of study. However, there are also likely to be cases of movement between domains being made out of necessity due to a lack of job opportunities in an individual’s domain of choice, for example.

2.1 Code and Data Availability

The Jupyter Notebooks used to perform the analysis are available from [8]. International RSE data can be obtained from [24], apart from some gender data which have not been publicly published at the time of writing.

3 Improving EDI in the research software community

In this section we investigate three areas that represent both challenges and opportunities for improving equity, diversity and inclusion (EDI) within the research software community. Through the material in this section, we hope to highlight possible reasons for the lack of diversity in the research software space while also providing some thoughts and guidance on how these can be addressed and how the community can work together to help improve a wide range of aspects of diversity. The material here is aimed at both developers of research software, and at team leaders, technical managers and individuals in other management and community leadership roles. For individuals who write software, we aim to offer thoughts and advice that can help improve understanding of the benefits of diversity and recognise situations that can contribute to a lack of diversity. For managers and team leaders, we hope to improve recognition of opportunities to address diversity challenges, both locally within teams and in the wider community in the context of events and activities.

3.1 Safety in similarity?

Diversity, whether in terms of gender, ethnicity, skills or other characteristics, is important in bringing different perspectives, ideas and experiences to a community. Diverse teams can lead to higher quality science [4] and improved technology business performance [22].

However, as highlighted by Merritt [21] in the context of hiring staff, individuals want to work with people who are like them and Lang & Liu [19] point out contradictory evidence for the benefits of diversity within teams. They highlight work by Byrne in the 1971 book “*The Attraction Paradigm*” that suggests certain similarities within teams can be helpful in ensuring effective working environments. However, they do also highlight a series of other work that supports the idea of diversity within teams and groups being beneficial. Van der Zee, et al. [31] highlight previous work that suggests that we tend to have a positive response to similarity while the opposite is true for dissimilarity. They also point out other previous work suggesting that we have an attraction to people who share similar attitudes and values to our own since this makes communication with them easier.

A very large study of over 20 years of scientific papers [12] shows that paper co-authors are more likely to be of similar ethnicity. However, where papers are the result of collaborations between authors at different locations, it was shown that this can result in them being published in journals with high impact factors and receiving more citations.

The above examples demonstrate that there is a perceived “safety” in being around people who have similar interests, backgrounds and/or values. We see that this is likely to work both ways and that in the case of a community, such as the RSE community or groups of software practitioners more generally, individuals choose whether or not to engage partly based on whether they see people like themselves within that community. Of course, while individuals may seek out others who they feel are similar or who they have things in common with, diversity can be extremely important and hugely valuable in providing different views, ideas, attitudes and perspectives. Ultimately this can lead to important benefits, even if there may be learning experiences to be had along the way. These include, for example, getting more used to working with people who may not look at every opportunity, challenge or research problem in the same way. Many of us will have had experiences that support this when attending community events and workshops which often involve breakout group discussions or problem-solving tasks.

Awareness is an important step here and this begins with **educating team leaders, community managers and event organisers to be more aware of situations where there is a lack of diversity and the benefits of addressing this**. Developing a stronger understanding of the benefits of diversity, but also the challenges that can arise in ensuring that diverse groups can collaborate effectively, is an important step towards addressing the lack of diversity in the scientific software community. This is especially important for individuals in leadership or organisational roles. Various guides, for example the Hopper

Conference Diversity Guide [28] and NumFOCUS DISCOVER Cookbook [23], provide detailed information on approaches for helping to ensure diversity at conferences and events. There is also a role for RSE group leaders in encouraging and enabling the RSEs in their teams to collaborate with other institutions, as this diversity might lead to higher impact of their work.

3.2 Increasing equity, diversity and inclusion at events

The events that we consider in this section include everything from large conferences and workshops to small local community events with only a small number of attendees. Large events often have dedicated organising teams and potentially co-chairs specifically dedicated to areas such as EDI. This is unlikely to be the case for a small community event with perhaps only a few 10s of participants. Nonetheless, being able to widen participation at events requires you to start with an understanding of the community the event is targeted at and an awareness of the particular aspects of diversity that you aim to improve [9]. One suggestion for improving equity and inclusion at conferences, from work looking at 30 conferences in the conservation and ecology domain, is to write up and promote details of the actions and processes followed to support improving EDI [29]. This work demonstrates that concerns around ensuring diversity at conferences and events are not specific to the research software community. The sort of open approach espoused through supporting and promoting EDI activities could be beneficial across many, if not all, domains. It would help to offer a demonstration of a conference’s efforts to support aspects of EDI and also provide evidence for the wider community to help identify what works and what is less successful. This should help to avoid repeating less successful approaches to addressing diversity and inclusion concerns across different domains.

Ensuring that event committees, speakers and panel members reflect the diversity that organisers would like to see among an event’s attendees has been effective in some fields. This builds on some of the ideas highlighted in Section 3.1. Research that looked at presenters across 21 meetings of the American Association of Physical Anthropologists found events that had either female, or both female and male organisers, resulted in a much higher percentage of women as first authors of presented papers and posters [18]. However a study by Bano and Zowghi of six software engineering conference series [2] found that having a female conference chair or program committee chair *did not* significantly affect the number of keynote speakers that were female, or the make up of the program committee. This work did not, however, look at all conference speakers.

While the figures shown in Section 2 show a lack of diversity within the RSE community, there is a perception, when attending RSE workshops and events, of gender and ethnic diversity being significantly better than the headline survey figures suggest. RSE events generally have a significant number of female speakers, workshop organisers and community leaders. We see with the Collaborations Workshop series, an event that focuses on general research software practice, a positive change in the diversity balance. It has had a policy of ensuring diversity of keynote speakers, and in 2020, 55% of steering committee members were

female, 42% of all speakers were female, 34% of attendees reported their gender as female, and 20% of attendees reported their ethnicity as non-white or mixed.

Nonetheless, the challenge remains of attracting more individuals from a wider range of backgrounds to get involved with RSE. A concrete action that can be taken here is to work to **increase the diversity amongst organisers, speakers and sponsors at RSE events and be more open about the approaches taken to support this**. The latter part of this recommendation should help organisers of other events, both within the research software community and beyond, to learn from and build on efforts to improve EDI. Ultimately this should help to accelerate the process of improving diversity throughout the research community.

3.3 An inclusive culture and safe space

An inclusive culture is an important foundation of diversity [6]. Research teams that include RSEs frequently have only one or two individuals in RSE roles. Having valuable but different skills, and different career aims, can make RSEs feel like the “odd one out” within a research group or team. As a result, RSE communities of practice have developed in recent years to support individuals in this field. Possibly due to previous perceived marginalisation, they tend to be open, welcoming communities. These communities provide a great opportunity to meet, network and collaborate with others who understand the challenges RSEs face in their day-to-day work.

At conferences and events, one way of formally promoting inclusion is having a clear and well-publicised *Code of Conduct*. The importance of a Code of Conduct and the challenges it can help to address and provide guidance on are highlighted in [10]. While still not commonplace at traditional academic conferences [11], it is unusual to find an RSE conference or workshop that does not have a Code of Conduct – learning from the experience of the open source community. To ensure best practice and clarity of the message provided, many events choose to build on widely accepted, open codes of conduct such as the template provided by the Geek Feminism wiki [13], and also have a diversity statement.

In the context of workplaces and research teams, it is also important that individuals feel that they fit in, both within the workplace environment itself and with the people based there. Cheryan et al. [7] call this “ambient belonging”. They undertook a series of studies within a computer science context looking at gender-based perspectives on the influence of different environments. This work showed that environments that had aspects that made them fit with computer science stereotypes reduced the interest of women participating in the field. Changing the environments to make them appear less like something that would be associated with computer science increased interest in the field. This provides an example of the sort of potentially small, but nonetheless significant, changes that can help to develop an inclusive culture, a welcoming environment and, ultimately, help to improve diversity. We consider that there may also be opportunities to help improve other aspects of diversity through identifying and

making similarly small environmental or organisational changes in different contexts, for example to improve the experience for people with disabilities.

The wide ranging use of codes of conduct at research software events and the recognition of the importance of communities in helping to provide individuals with a place where they fit in is significant. It suggests that the research software community is developing events that are explicit in their desire to be open and diverse, with clear statements on acceptable behaviour. We see that progress is being made in developing an inclusive culture and providing environments through events and activities that potential participants feel are open and welcoming. It is of vital importance that this continues. There are two concrete recommendations from this analysis. Firstly, **within RSE groups, both team leaders and members should be aware of the importance of their working environment and how potentially small environmental changes may affect other team members' feelings of inclusion and belonging.** Secondly, in the context of community events and activities, **event organisers should look to highlight their support for diversity and inclusion through the use of a diversity statement and provide a clear code of conduct highlighting acceptable behaviour.**

The Software Sustainability Institute runs a fellowship programme recognising the diverse roles and skills of those working to promote research software practice. An evaluation of the programme showed it plays an important role in supporting communities of best practice and skills transfer, and that a significant benefit is the way it has raised the profile of software in research, and those people who develop and advocate for it [27]. This has had positive effects for those who may previously have considered themselves as ‘outsiders’ in the role, or lacked confidence. This is exemplified by a comment from a female respondent:

“Despite getting a PhD partially from a computer science programme, I could see that my skills and knowledge were always at least to some extent dismissed or doubted [...] since being elected a SSI fellow I most definitely observed a significant drop in mansplaining... I have little doubt that the SSI fellowship was a significant [reason] I got my current position as (Head of Division at a Supercomputing Center).” [27]

4 Discussion and Conclusions

Our research has shown that the gender diversity of RSEs is similar to the field of software engineering, but does not reflect the gender balance of the academic research workplace. This is particularly true where RSEs are working in domains such as the biosciences, geosciences and medicine. Ethnic diversity among RSEs is worse than in the wider software industry, but it is unclear why. The number of RSEs with disabilities reflects the academic research workplace, and is poor in comparison to the IT profession. Given that 19% of the UK working age population report a disability, there is much to be done to provide a equitable workplace for RSEs with disabilities. Differences in workplace culture, environment or incentives may be factors, and further research is required. However,

general interventions to improve diversity appear to be increasing gender and ethnic diversity at events.

We have highlighted the importance of communication and collaboration between individuals building software to support research and the computational scientists and researchers that they collaborate with. We have also shown that diversity among all parties in these collaborations can lead to better communication, a wider range of ideas and perspectives, and, ultimately more effective collaborations that produce higher quality outputs. The three areas discussed in Section 3 as key opportunities for improving equity, diversity and inclusion within the field of research software can be summarised as follows:

- **Safety in similarity?** The perceived safety of looking to collaborate and work with individuals who have similar interests, backgrounds or values is widely recognised. While there is research that suggests there can be benefits to this approach, there is also extensive work highlighting the benefits of diversity. However, this is something that can be approached from two different perspectives and individuals may be more likely to join and engage with a community if they see people like themselves within that community.
- **Increasing diversity at events:** It is important to highlight and promote actions being taken to improve EDI in the context of events. It is also important to ensure that both organising groups and speakers reflect the level of diversity that an event’s organisers would like to see among its participants.
- **An inclusive culture and safe space:** This is a very important aspect of both participating in a community, and being part of an RSE group, for everyone involved. A well-publicised Code of Conduct and diversity statement with clearly defined processes supporting them are key elements in helping to ensure this within the context of events. In RSE groups, leaders and team members need to be aware of the importance of an inclusive culture and working environment.

Poor diversity in the community of research software developers is likely a result of the low levels of diversity in the disciplines that currently form the major path towards an RSE role or career. We have identified scope to widen the range of areas from which RSEs are recruited and it is hoped that this can be achieved by better use of language in the way that RSE roles are advertised and in the way that RSE is promoted, more generally. This may also have relevance for professional software practitioners looking to get involved in the research community, as it is clear that there are many people with “non-traditional” degrees seeking careers developing software. There is a role for professional bodies, such as the British Computer Society, Association for Computing Machinery and IEEE Computer Society, to support and embrace these career paths and help improve the gender diversity at undergraduate and postgraduate level.

Nonetheless, we feel that despite the knowledge gained from our extended analysis of existing survey data, and the pre-existing material that we have referenced, this is a hugely complex but very important area. It would benefit extensively from additional evidence that could be gathered through a range

of further empirical studies. Some of the areas that may be considered most important for such studies in the short term include:

- Exploring why levels of diversity among RSEs are lower than in many of the areas of study and research that already feed into RSE careers, such as Computer Science and Physics and Astronomy. What are the levels of diversity in other career paths for these subjects? What influences the career choices that individuals make and are there specific aspects that steer them away from an RSE career?
- Trialing approaches for increasing diversity and inclusion at events, including workshops and conferences. Gathering statistics on the relative success of these approaches and their contribution to improving diversity within the computational science developer community.
- Looking at opportunities to further increase “domain mobility” through taking advantage of existing courses/training material or offering new information that can help individuals move more practically between working in different domains. Look at take-up of such opportunities across different domains and use this as a basis for longer-term analysis on how diversity in the RSE community changes over time.

Once individuals become RSEs and engage with the RSE community, the RSE community meets many of the requirements to ensure that it welcomes and supports diversity. The open, inclusive, nature of the community is of great importance here and, using gender diversity as an example, there are already several women in highly-visible leadership roles.

Developing software for science and the wider computational research domain is challenging. The relatively new field of research software engineering encompasses many of the individuals undertaking these software development tasks, regardless of whether their official job title considers them to be a researcher, professor, software engineer or RSE. It is clear that research software engineering, and research software in general, shares many diversity challenges with the wider software development field but we feel there are many opportunities to address this. This begins with increasing awareness of the different aspects raised in this paper and using this as a basis to develop the environment and opportunities to help build equity, diversity and inclusion within the community. We hope that the research software community continues to grow, and become more diverse by sharing and learning from other practitioners.

5 Acknowledgement

We thank Simon Hettrick, James Graham and Rob Haines for their input, and Olivier Philippe who originally processed the international RSE survey data. NCH and CJ acknowledge support from EPSRC / BBSRC / ESRC / NERC / AHRC / STFC / MRC grant EP/S021779/1 for the UK Software Sustainability Institute. JC acknowledges support from EPSRC grant EP/R025460/1.

References

1. Apa, A.: Driving Gender Equality in the Tech Industry: Breaking down unconscious bias (Jun 2018), [Online] Available at: <https://womeninhpc.org/hpc/driving-gender-equality>
2. Bano, M., Zowghi, D.: Gender disparity in the governance of software engineering conferences. 2019 IEEE/ACM 2nd International Workshop on Gender Equality in Software Engineering (GE) (May 2019). <https://doi.org/10.1109/ge.2019.00016>
3. Brett, A., Croucher, M., Haines, R., Hettrick, S., Hetherington, J., Stillwell, M., Wyatt, C.: Research Software Engineers: State of the Nation Report 2017 (April 2017). <https://doi.org/10.5281/zenodo.495360>
4. Campbell, L.G., Mehtani, S., Dozier, M.E., Rinehart, J.: Gender-heterogeneous working groups produce higher quality science. PLoS ONE **8**(10), e79147 (Oct 2013). <https://doi.org/10.1371/journal.pone.0079147>
5. Capretz, L.F., Ahmed, F.: Why do we need personality diversity in software engineering? ACM SIGSOFT Software Engineering Notes **35**(2), 1–11 (Mar 2010). <https://doi.org/10.1145/1734103.1734111>
6. Chamorro-Premuzic, T.: How to design a diversity intervention that actually works (Jun 2019), [Online] Available at: <https://www.forbes.com/sites/tomaspremuzic/2019/06/20/how-to-design-a-diversity-intervention-that-actually-works/>
7. Cheryan, S., Plaut, V.C., Davies, P.G., Steele, C.M.: Ambient belonging: How stereotypical cues impact gender participation in computer science. Journal of Personality and Social Psychology **97**(6), 1045–1060 (2009). <https://doi.org/10.1037/a0016239>
8. Chue Hong, N., Cohen, J., Jay, C.: RSE Diversity Analysis Notebooks (Apr 2021). <https://doi.org/10.5281/zenodo.4662166>
9. Collis, T.: Improving diversity at hpc conferences and events (2016), J. West, L. Rivera and N. Chue Hong (eds.). [Online] Available at <http://www.hpc-diversity.ac.uk/best-practice-guide/access-best-practice>
10. Favaro, B., Oester, S., Cigliano, J.A., Cornick, L.A., Hind, E.J., Parsons, E.C.M., Woodbury, T.J.: Your science conference should have a code of conduct. Frontiers in Marine Science **3** (Jun 2016). <https://doi.org/10.3389/fmars.2016.00103>
11. Foxx, A.J., Barak, R.S., Lichtenberger, T.M., Richardson, L.K., Rodgers, A.J., Webb Williams, E.: Evaluating the prevalence and quality of conference codes of conduct. Proceedings of the National Academy of Sciences **116**(30), 14931–14936 (Jul 2019). <https://doi.org/10.1073/pnas.1819409116>
12. Freeman, R.B., Huang, W.: Collaborating with People Like Me: Ethnic Coauthorship within the United States. Journal of Labor Economics **33**(S1), S289–S318 (Jul 2015). <https://doi.org/10.1086/678973>
13. Gardiner, M., Aurora, V., Smith, S., Benjamin, D.: Geek Feminism Wiki: Conference anti-harassment/Policy, [Online] Available at: https://geekfeminism.wikia.org/wiki/Conference_anti-harassment/Policy, Accessed 10th Aug 2020.
14. Graham, H., Raeside, R., Maclean, G.: Understanding the status of under-represented groups in the information and communication technologies: A report to the engineering and physical sciences research council (July 2017), [Online] Available at <https://epsrc.ukri.org/newsevents/pubs/napierdiversityreport/>
15. HESA: HE staff by activity standard occupational classification, atypical marker, contract marker, mode of employment, terms of employment, contract levels, sex, age group, disability, ethnicity and academic year, [Online] Available at: <https://www.hesa.ac.uk/data-and-analysis/staff/table-3>

16. HESA: HE student enrolments by subject area and sex, [Online] Available at: <https://www.hesa.ac.uk/data-and-analysis/sfr247/figure-14>
17. Institute of Physics: What does a physicist look like? (Mar 2016), [Online] Available at: http://www.iop.org/publications/iop/2016/file_67244.pdf.
18. Isbell, L.A., Young, T.P., Harcourt, A.H.: Stag parties linger: Continued gender bias in a female-rich scientific discipline. *PLoS ONE* **7**(11), e49682 (Nov 2012). <https://doi.org/10.1371/journal.pone.0049682>
19. Liang, T., Liu, C., Lin, T., Lin, B.: Effect of team diversity on software project performance. *Industrial Management & Data Systems* **107**(5), 636–653 (May 2007). <https://doi.org/10.1108/02635570710750408>
20. Mathieson, S.A.: How diversity spurs creativity in software development (Dec 2019), [Online] Available at: <https://www.computerweekly.com/feature/How-diversity-spurs-creativity-in-software-development>. Accessed 26th Jan 2021.
21. Merritt, C.: Why do we keep hiring ourselves? (Sep 2018), [Online] Available at: <https://womeninhpc.org/diversity-and-inclusion/why-do-we-keep-hiring-ourselves>
22. NCWIT: What is the impact of gender diversity on technology business performance? (2014), [Online] Available at: https://www.ncwit.org/sites/default/files/resources/impactgenderdiversitytechbusinessperformance_print.pdf.
23. NumFOCUS: NumFOCUS DISCOVER Cookbook (Diverse & Inclusive Spaces and Conferences: Overall Vision and Essential Resources) (2019), [Online] Available at: <https://discover-cookbook.numfocus.org/>. Accessed 13th Aug 2020.
24. O. Philippe et al.: softwaresaved/international-survey: Public release for 2018 results (Mar 2019). <https://doi.org/10.5281/ZENODO.2585783>
25. Society, B.C.: BCS Diversity Report 2020: ONS Analysis (Jun 2020), [Online] Available at: <https://www.bcs.org/media/5766/diversity-report-2020-part2.pdf>.
26. Stahl, G.K., Maznevski, M.L., Voigt, A., Jonsen, K.: Unraveling the effects of cultural diversity in teams: A meta-analysis of research on multicultural work groups. *Journal of International Business Studies* **41**(4), 690–709 (Nov 2009). <https://doi.org/10.1057/jibs.2009.85>
27. Sufi, S., Jay, C.: Raising the status of software in research: A survey-based evaluation of the software sustainability institute fellowship programme [version 1; peer review: 3 approved with reservations]. *F1000Research* **7**(1599) (2018). <https://doi.org/10.12688/f1000research.16231.1>
28. The Hopper Fund: The Hopper Conference Diversity Guide: Do Better at Conference Diversity, [Online] Available at: <https://conference.hopper.org.nz/>. Accessed 13th Aug 2020.
29. Tulloch, A.I.T.: Improving sex and gender identity equity and inclusion at conservation and ecology conferences. *Nature Ecology & Evolution* **4**(10), 1311–1320 (Aug 2020). <https://doi.org/10.1038/s41559-020-1255-x>
30. Universities UK: Patterns and trends in UK higher education 2018 (Sep 2018), [Online] Available at: <https://www.universitiesuk.ac.uk/facts-and-stats/data-and-analysis/Documents/patterns-and-trends-in-uk-higher-education-2018.pdf>.
31. Van Der Zee, K., Atsma, N., Brodbeck, F.: The influence of social identity and personality on outcomes of cultural diversity in teams. *Journal of Cross-Cultural Psychology* **35**(3), 283–303 (May 2004). <https://doi.org/10.1177/0022022104264123>